DOCUMENT RESUME

ED 302 830 CS 009 481

AUTHOR McKenna, Marian J.
TITLE The Development and Validation of a Model for Text Coherency.
PUB DATE Nov 88
NOTE 18p.; Paper presented at the Annual Meeting of the National Reading Conference (38th, Tucson, AZ, November 29-December 3, 1988).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Coherence; Discourse Analysis; Higher Education; Holistic Evaluation; Models; Reading Writing Relationship; *Text Structure; Writing Evaluation; Writing Research
IDENTIFIERS *Text Coherence; *Text Factors

ABSTRACT

To discover what variables are involved in the production of text coherence, and how cohesion and coherence are different, a study collected 30 papers randomly selected from over 200 papers written by incoming freshmen college students at a local community college. The papers were rated by 21 state conference participants. Raters were in-service language arts or reading teachers, with teaching experience ranging from 5 to 20 years. All raters had participated in a Colorado Writing Assessment Project, and had prior training and experience with holistic rating of essays. Each rater received a packet including a copy of one of seven scales, 30 papers for rating, and training notes. Each of the 30 papers was read three times for each of the seven scales. Bamberg's scale for coherency was the dependent variable. Six scales, the independent variables of coherence, consisted of: cohesion; context; focus; grammar; intent; and structure. Results indicated interdependency between intent, focus, structure, and context. Intent was a significantly stronger predictor of text coherency than was cohesion or grammar. Results also indicated that if the variables of intent, structure, focus, and cohesion were highly present in a text, then a lack of grammar and context did not greatly affect the text's coherency level. It was further demonstrated that an author's apparent intent in writing a text was a strong predictor for the coherency of the text. (Three tables of data are included.) (MM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED302830

Marian J. McKenna
University of Colorado
4818 Mac Intosh Place
Boulder, CO 80301
(303) 444-4920

THE DEVELOPMENT AND VALIDATION OF A MODEL FOR TEXT COHERENCY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Marian McKenna
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Mc Kenna
Running head: Model for Text Coherency

CS009481

THE DEVELOPMENT AND VALIDATION OF A MODEL FOR TEXT COHERENCY

If it is true that information in the world is doubling at a rate of every 22 months, then more than ever, educators must be synthesizers of information across many fields of inquiry in order to be able to extract it from the research and to impart it to their students. An area of inquiry where this explosion of information is most apparent is in the field of writing. The increased number of studies being done in Discourse Analysis demands that a clear distinction in meaning be made between the two terms of cohesion and coherence. If this is not done, research that could be useful, specifically to composition and reading teachers, will never be accessible to them. It is felt that the concepts of cohesion and coherence are central to the issue of making meaning in a language interchange. Children can be taught to use transitional phrases in their writing and to be guided by such transitional phrases in their reading. The research work being done in discourse analysis has aided students and teachers as it has enabled them to incorporate the use of higher order thinking skills in the classroom. The work of Meyer (1975, 1984) in structure of text and its effect on recall has indicated that students who are instructed in the structure of a text are more successful in the recall of that text and are able to produce text having a similar structure. Halliday and Hasan (1976) investigated the property of cohesion in text and defined

it as that device which allows a text to flow in a connected way from sentence to sentence and paragraph to paragraph. This was followed by a study done by Witte and Faigley (1981) that indicated that the type and number of cohesive devices, such as pronomial references and collocation that are used in a text, are indicators of the quality of the essay.

Bamberg (1984) conducted a reanalysis of the description papers used in the NAEP study (1969-1974), using a holistic scale that she developed to look at the property of coherency as a function of a whole text rather than as individual paragraphs. Bamberg assessed coherence in terms of a list of features that work together to create coherence in a whole text. Bamberg's study went a long way towards advancing the idea of coherency as a distinct property of an entire text. However, Bamberg's scale does not go far enough in delineating the specific features of coherence, nor the importance of any one of those features as a determinant of coherence. Further, Bamberg's coherency scale covers so many aspects of text that it proved unwieldy as a research tool and overwhelming for use in the classroom by teachers and beginning or unskilled writers. For example, any one of the five ratings (4-0) that can be received on Bamberg's scale, cover the textual aspects of topic, focus, context, structure, cohesion and grammar. It was noted in conducting the current study, that the raters using Bamberg's Coherency scale took the longest time to rate the 30 papers. Therefore, Bamberg's work brought us closer to an understanding of the nature of coherence as it manifested in text, but

not to the point where it was useful as a tool for diagnosis, instruction and self-assessment for students and teachers in the classroom.

The purpose of this study then was two-fold. The initial intent was to clearly establish the idea that there is a difference between the textual properties of cohesion and coherence. The two terms are often used interchangeably in the literature, rendering discourse analysis studies less accessible than they could be, especially to educators. In this study, Halliday and Hasan's definition of cohesion, as that property of text whereby referential devices are used to create a connected discourse, was accepted.

If this is cohesion, then the second goal of this study was to determine a definition for coherency in English texts. Coherency is a result of the degree of the interaction between a reader and a writer. This study looked only at those aspects of coherency which can be discerned from a textual point of view. Six constructs were outlined and defined as being the variables that will determine the presence of coherence in an English text. They are: Author's Apparent Intent (hereafter referred to as Intent), Context, Focus, Structure, Cohesion and Grammar. In breaking out these six features of text, it is believed that we can not only talk about whether a paper is coherent, we can also talk about why it is or is not coherent.

Method

The first premise guiding this study is that coherence is something different from cohesion. Following this is the second

premise, that cohesion is a text-bound property of the language as defined by Halliday and Hasan. On the other hand, coherence encompasses both text and a world knowledge base of the text user. That is, meaning is mediated between two language users based on the amount of "conversation" that they share.

This study demonstrated the soundness of these two premises by presenting a search of the literature in the fields of psychology and linguistics for a working definition of coherence, to discover what variables are involved in the production of coherence, and then applying them to a series of college students' compositions. From the literature and from student papers, variables in a text that contribute to the presence of coherence were discovered. The variables of coherence, one of which is cohesion, were defined and developed into holistic scales. The remaining variables contributing towards coherence are Grammar, Structure, Focus, Context, and Intent. Raters then used holistic scales developed for these variables to rate a series of compositions for the presence of any particular variable. These variables were chosen for use after rigorous study of many student essays, review of the current literature and a pilot study to determine the clarity of each of the scales used to measure the extent of the presence of each of these variables in a text. Each scale developed for this study was calibrated to match Bamberg's scale for measurement purposes, but each scale was constructed to look only at the specific, designated variable. Results from these ratings were correlated and studied to determine validity and the importance of the

individual variables in the development of coherence in an English text and how the understanding of these variables might be useful in a pedagogical setting.

Sample

The 30 papers used for this study were randomly selected from over 200 papers that were written by incoming freshmen college students at a local community college. The students wrote these papers extemporaneously on the first day of class. These papers were then used to determine to which composition class the students would be assigned.

Procedures

The actual gathering of the data for this study was done in a single, two-hour period, during a state conference. There were 21 participants and three facilitators for the study. The participants or raters for this study were all in-service language arts or reading teachers, with teaching experience ranging from five years to 20 years. All of these raters had participated in a Colorado Writing Assessment Project under the auspices of the Colorado State Department of Education, and had prior training and experience with holistic rating of essays.

Each rater in the study received a packet containing a copy of one of the seven scales, 30 papers to be rated using that scale, prototypes for at least four of the levels on the scale, and training notes particular to that scale. In this way, each of the 30 papers were read

three times for each of the seven scales. Bamberg's scale for coherency was the dependent variable. The six scales developed for this study as the variables of coherence were the independent variables. The following is a definition of each one of the six independent variables:

Cohesion--That property of text whereby the author uses such lexical ties as reference, substitution, ellipsis, conjunction and lexical reiteration and collocation to create connections which tie the discourse together as a unit.

Context--The physical and social setting which the author uses to carry out the communication goal. The physical setting refers to the determination of the time and place within which the text can be placed. The social setting of the text refers to the personal setting of the communication endeavor, such as whether or not the communique is a business one, an academic one or a personal communication.

Focus--The implied contract between the reader and the writer that the writer will address one topic only and will remain true to that topic throughout the text.

Grammar--Those aspects of text which are concerned with traditional grammar usage as it is taught in language arts or English classrooms (i.e., Warriner's Grammar). Mechanics such as spelling, punctuation, sentence structure and word usage are all included here.

Intent--The concern here is with the author's apparent intent and the clarity of that intent as it was conveyed by the author's message in the text. For example, in a given text, is it clear that the

author's intent is to convince the audience of one side or another of an issue?

Structure--The organizational plan that the author uses to fulfill the communication goal. For example, an author may choose to use a comparison and contrast paradigm to reach a certain textual and communication goal.

Results

The questions posed in this study lent themselves to a correlational analysis to see how the independent variables of intent, focus, context, structure, cohesion and grammar are related to the dependent variable of coherence and to each other. To answer this question, a Pearson Correlation Coefficient was computed. The results of these calculations were sufficient to address the first hypothesis that the independent variables correlate with the property of coherency in a text. Intent had the highest correlation with coherence, at .68, followed by context at .66, structure at .65, focus at .63, grammar at .46, and finally cohesion at .36. It is of interest to note that cohesion has the lowest correlation with coherence, which strengthens the claim of this study that coherence and cohesion are different properties of text. However, the independent variables of focus, structure, context and intent are highly correlated with each other. However, even the disattenuated correlation coefficients were not over .90, indicating that the raters were able to discriminate between these variables to some degree. In a purely logical sense, it would seem unlikely that we would find an example of text that was very high in

intent, yet very low in focus or structure. It seems safe to assume that there is some interdependency between the variables of intent, focus, structure and context.

A t-test was then conducted to test the statistical differences between correlations based on the same sample and to address the second question of whether or not some of the independent variables are more highly correlated with coherence than others. This test was calculated using the highest and the lowest correlations until the point was reached where the null hypothesis remained tenable. Table 1 reports the findings of these calculations.

Insert Table 1 about here

The reporting method used in this table is a common tabulation for reporting multiple comparisons. Where two variables are underscored by a common line, the null hypothesis remains tenable. The null hypothesis in this case states that there is no significant difference between the correlations of the independent variables to coherence. The results of this t-test indicate, however, that there is a statistical difference in how these variables correlate with coherence. Intent is a significantly stronger predictor of coherency in a text than is either cohesion or grammar. Further, context is a significantly stronger predictor of coherency in a text than is cohesion, as is structure to cohesion and focus to cohesion. Other correlations remain tenable.

The final question posed in this study is whether or not cohesion

correlates differentially with coherence than do the other independent variables. This hypothesis is proposed to strengthen the claim of this study, that cohesion and coherence are different properties of text and that in fact, cohesion is just one factor that contributes to coherency in a text. Again, a t-test was conducted to determine the statistical significance of this claim.

The results of this test indicate a different pattern of correlation of cohesion to the independent variables than that of coherence to the indpendent variables (see Table 2). The variables

---

Insert Table 2 about here

---

that correlate with coherence are different than those that correlate with cohesion. These results indicate that cohesion may not be a strong predictor of coherence in a text. Grammer, however, may be a strong predictor for the presence of cohesion in a text.

To illuminate these relationships even further, a Step-wise Forward Regression Analysis was conducted (see Table 3).

Insert Table 3 about here

The total variance accounted for in Table 3 by these four variables then is .56. The variables of grammar and context were not included in the regression equation. If we have an interrater reliability of 90%, as in the case of focus, then 10% of the variance could be construed as measurement error. In this case, 34% of the

variance is unaccounted for and could be speculated as being the result of relevant variables that have not been included in this study and may be somewhat contributing to the variance here. It makes sense that 34% of the variance is left unexplained, as the writers in this study wrote their essays for an academic purpose and context and the readers in this study were teachers and were familiar with that context. However, there was some variability in the age of the readers and in the length of their teaching experience as well as experience with holistic ratings. This finding is in keeping with the researcher's belief that coherency in a text is the result of an interaction between the reader and the writer of that text.

In the regression analysis, grammar and context were not included in the equation as they would not significantly have contributed any new variance to the equation. Grammar did not correlate very highly with the dependent variable of coherence or with any of the independent variables. Context, on the other hand, did correlate highly with coherence (.66). However, it was also highly correlated with the independent variable of intent (.70). We can assume here that the information contributed by context was redundant with that of intent.

Therefore, one of the first patterns for the model of coherence that emerges from this analysis is that if the variables of intent, structure, focus and cohesion are highly present in a given text, then a lack of grammar and context will not greatly affect the coherency level of that text. It was further demonstrated here that an author's apparent intent in writing a text is a strong predictor for the

coherency of that text.

Discussion

The title of this study is "The Development and Validation of a Model for Text Coherency." As a summation to the findings of this study, an operationalized definition of coherency from the textual side of the communications process is proposed: a high level of textual coherency is achieved when an author's apparent intent is clear to the reader, that intended message is supported by a strong structure in the text, there are no outstanding violations of the focus contract, and the text flows in a smooth and logical way from sentence to sentence and paragraph to paragraph, as measured by the holistic scales presented in this study.

Coherency of a text is a much more global function of a text than is cohesion or indeed any of the other independent variables included in this study. The independent variables discussed here need to be operating together to some degree to create coherency. The variables of particular importance in determining coherency are those of intent, structure, focus and cohesion.

Our concern as reading and writing educators should be with the communication endeavor as a whole. Our interest should be with how text is used and understood in a real world setting, making our students aware of their intent in both reading and writing as well as the needs of their audience, when they are the writers or themselves as an audience when they are reading. If we can teach basic readers and writers to consider such issues as intent and awareness of audience, we

will be teaching reading and writing as higher order thinking skills. This study, therefore, has the potential to have an impact for further research in discourse analysis as well as in the field of reading and writing education.

It is hoped that as we gain a more clear understanding of the nature of coherence and of discourse production, that we can take our students from being passive readers and writers, to being truly interactive readers and writers. That is, we can help our students to enter into the conversation that is civilization.

References

Bamberg, B. (1984). Assessing coherence: A reanalysis of essays written for the National Assessment of Educational Progress, 1969-1979. Research in the Teaching of English, 18(3), 305-319.

Glass, G. V., & Hopkins, K. D. (1984). Statistical methods in education and psychology (2nd ed.). New Jersey: Prentice-Hall.

Halliday, M. A. K., & Hasan, R. (1976). Cohesion in English. London: The Longman Group.

Meyer, B. J. F. (1975). The organization of prose and its effect on memory. Amsterdam: North-Holland Publishing Co.

Meyer, B. J. F., & Freedle, R. O. (1984, Spring). Effects of discourse type on recall. American Educational Research Journal, 21, 121-143.

Witte, S. P., & Faigley, L. (1981, May). Coherence, cohesion and writing quality. College Composition and Communication, 32, 189-204.

Table 1

Multiple Comparisons for Significance of Correlations of the Independent Variables to Coherence

H_0: $P_{31} = P_{32} \cdot \cdot \cdot P_{3i}$

Intent	Context	Structure	Focus	Grammar	Cohesion

Note: The null hypothesis is rejected for: Intent/Cohesion (t = 1.94); Intent/Grammar (t = 1.55); Context/Cohesion (t = 1.69); Structure/Cohesion (t = 1.64); Focus/Cohesion (t = 1.43).

$\alpha_1 = .10$.

Critical t = 1.31.

$p < .001$.

Table 2

Multiple Comparisons for Significance of Correlations of the Independent Variables and Coherence, to Cohesion

Grammar	Coherence	Intent	Structure	Context	Focus

Note: The null hypothesis is rejected for: Grammar/Focus (t = 1.96); Grammar/Context (t = 1.53); Grammar/Structure (t = 1.64); Coherence/Focus (t = 1.84).

α_1 = .10.

Critical t = 1.31.

$p < .001$.

Table 3

Step-wise Forward Regression Analysis of the Correlations of the Independent Variables to Coherence

Step No.	Variable Entered	Multiple R	Beta	Adjusted R Squared
1	Intent	.6788	.167	.4414
2	Structure	.7364	.344	.5084
3	Focus	.7588	.319	.5269
4	Cohesion	.7863	.220	.5573